International Journal for Housing Science and Its Applications
Publish August 6, 2025. Volume 46, Issue 3 Pages 7083-7090

https://doi.org/10.70517/ijhsa463608

Analyzing the Artistic Characteristics of Guanzhong Drum and Drum Dance and the Construction of Diversified Inheritance Paths Using Computer Vision Technology

Sitong Chen[1] and Jian Yang[1,*]
[1] College of Literature and Media, Xi'an FanYi University, Xi'an, Shaanxi, 710105, China
Corresponding authors: (e-mail: sachi12345@163.com).

Abstract The development of computer vision technology provides more feasibility for the inheritance and development of Guanzhong Drum and Gong Dance, a traditional folk dance. This paper analyzes the dance movements of Guanzhong Drum and Drum Dance, summarizes the artistic characteristics embedded in the dance movements, and assists in the construction of the inheritance path of the dance as the research idea. The basic features of Guanzhong Drum and Drum Dance are briefly analyzed to pave the theoretical framework of the study. Afterwards, the movement characteristics of Guanzhong Drum and Drum Dance are converted by using the universal format BVH, and the pre-processing of the dance movement data is carried out. At the same time, 3D CNNs are used as the action recognition algorithm, and the ReLU function is used as the activation function to prevent the model from overfitting, so as to propose a Guanzhong Gong and Drum Dance action recognition model based on 3D CNNs. Compared with similar model algorithms, the proposed model algorithm can recognize the Guanzhong gong dance with an accuracy of 85.00% and above, which shows a superior performance in action recognition.

Index Terms Guanzhong gong dance heritage, dance movement recognition model, 3D CNNs, ReLU function

I. Introduction

In today's rapidly developing society, people are gradually forgetting the richest cultural heritage left by their predecessors. Fully exploring the cultural value embedded in this important intangible cultural heritage, especially the artistic and communication value therein, is essential for better inheriting and promoting Chinese culture [1].

Guanzhong region of Shaanxi has a long history and rich cultural heritage. It is the ancient capital of the pre-Qin emperors and has a rich cultural heritage, and in terms of dance, the folk dances in Guanzhong region of Shaanxi can also be described as varied and diverse [2], [3]. Among them, Guanzhong gong and drum dance has a strong cultural atmosphere and rich mass base in individual villages and towns of local folklore, but its popularity is not high, there are individual “village fever” phenomenon, government public information shows that the gong and drum dance score inheritor is less than 20, and there is a serious simplification phenomenon in the dance movement. This kind of dynamic non-genetic inheritance can not only rely on personal physical memory and simple image records, and need to combine modern technology to deeply analyze the artistic characteristics of dynamic non-legacy, and build a suitable path for its cultural inheritance [4], [5].

With the continuous development of science and technology, computer vision technology has gradually become an important tool and means in cultural heritage protection. Using computers and digital image processing algorithms, computer vision technology can extract information about cultural heritage through the analysis and processing of images and videos to realize the protection, restoration and inheritance of cultural heritage [6]-[8]. In cultural heritage protection, computer vision technology can be applied in a variety of ways, including gesture capture, property rights of cultural relics, three-dimensional virtual exhibitions, cultural relics detection and tracking, etc., which breaks through the limitations of the traditional image recording from the multi-dimensional aspects of time and space, culture, and dissemination, to improve the level of protection of cultural heritage, to strengthen the cultural exchanges and promotion, and to achieve the sustainable development of cultural heritage [9]-[12].

In this paper, we firstly briefly summarize the dance categories of Guanzhong Drum and Drum Dance and the characteristics of the classical repertoire, which serves as the theoretical basis for the research and design of this paper. Secondly, it discusses in detail the calculation method of 3D coordinates of nodes of Guanzhong Drum and Drum Dance in the form of universal format BVH, so as to propose the pre-processing method of motion capture data. The 3D CNNs algorithm is chosen as the motion recognition algorithm, and the computation method of the dance movement features and the overfitting phenomenon prevention method are designed to build the Guanzhong gong dance motion recognition model based on the 3D CNNs algorithm. Finally, the proposed model is used to segment the action sequences of Guanzhong Gong and Drum Dance in the dataset, and to test the effect of human displacement localization in the dance and evaluate the accuracy of dance action recognition.

II. Artistic Characteristics of Guanzhong Drum and Gong Dance

The splendid and colorful Shaanxi folk dances, promoted by the splendid national culture and the ancient music and dances of the Zhou, Qin, Han and Tang dynasties, the hundred operas as well as the new rice-planting song movement, constitute the precious remnants of Shaanxi's dance culture, and Shaanxi can be said to be one of the representative and important regions of China's Han folk dances. In recent years, more than two hundred and sixty kinds of folk dances have been successively excavated in the province. According to the purpose and function of Shaanxi folk dance activities, as well as its artistic form and characteristic analysis, it can be basically summarized into three major categories, i.e., social fires, folk rituals, and folk plays. Folk fire is the main body of Shaanxi folk dance, which is rich in content and diverse forms, of which the rice-planting songs and drums are the most widely represented. In northern Shaanxi Province, rice-planting songs are the mainstay of folk dances, and various folk dances such as waist drums and drum whips are incorporated. In Guanzhong, social fire is the main form of performance, with strong gong and drum performances and comprehensive social fire dances. The Guanzhong Drum and Drum Dance has different requirements on props for the interpretation of different drums and drums in the performance. For example, the “Cow Pulling Drums” focuses on the interpretation of the drummer. The “Mighty Drums of War” is an interactive performance of the ceremonial snare drum squad. Different interpretations of the Guanzhong Drum and Drum Dance have different choreographies for the props.

The dance roles of the Guanzhong Drum and Drum Dance are divided into: the drummer, the flag head, the ceremonial drummers, and the accompaniment team (gongs and raos). The Guanzhong Drum and Drum Dance takes many different forms, and each form has its own requirements for role assignments and number of dancers. The traditional form handed down during the Republic of China period is the Daqi Dance, in which a big drum and a flag head are accompanied by dozens of small drums and dozens of windings, with an equal number of small drums and raos. Nowadays, the Guanzhong Drum and Drum Dance is very flexible in the use of props, and there are some that cater to stage performances by removing the flag head and the drums. There are also the accompaniment team and the ceremonial drums that correspond to each other and reflect each other. It can be said that the form is ever-changing, novelty.

III. Guanzhong gong dance action recognition model based on 3D CNNs

III. A. Dance Motion Capture Data Preprocessing

In this paper, motion capture data in the form of generic format BVH is used. The BVH file contains the definition of the human skeleton model and the node motion information recorded in the form of Euler angles. When analyzing the human body motion, Euler angles only describe the relative angular relationship between nodes and cannot visually represent the relative positions between nodes. In order to clarify the spatial position of each node in the skeleton model for analysis, the motion data recorded in BVH format is converted to 3D coordinate form in the preprocessing stage. The human skeleton model of motion capture data in this paper contains 26 nodes see Fig. 1.

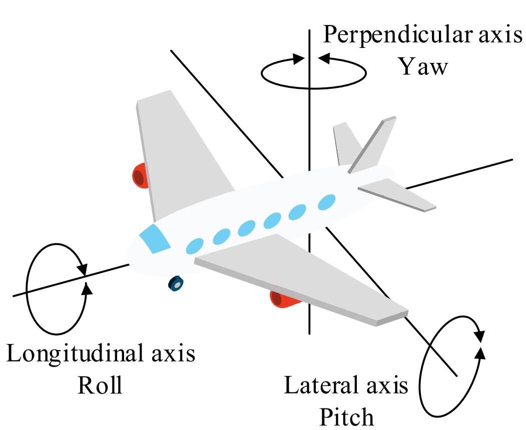

Figure 1: Rotation diagram of“Roll-Pitch-Yaw"

In the human skeleton model, for a child node J and its parent node J_p, the Euler angle records the angle of rotation of that parent node J_p along the X, Y and Z axes. After J_p is rotated by a certain angle, its angle of rotation along the X, Y, and Z axes changes, creating a new angular displacement. In the Cartesian coordinate system, it is agreed that the angles of rotation of J_p around the ZXZ-axis according to the "Roll-Pitch-Yaw" setting are, in order, the Roll angle r, the Pitch angle p and the Yaw angle y. Assuming that the relative positions of J and J_p at the beginning, i.e., the initial offset of J with respect to J_p is (x_0, y_0, z_0), and the three-dimensional coordinates of J_p are (x_{p0}, y_{p0}, z_{p0}), the three-dimensional coordinates of J, (x_{J0}, y_{J0}, z_{J0}), can be obtained by adding its initial relative position with J_p's initial relative position and J_p's 3D coordinates to obtain equation (1):

$$(x_{J0}, y_{J0}, z_{J0}) = (x_0, y_0, z_0) + (x_{p0}, y_{p0}, z_{p0}) \tag{1}$$

After a certain motion occurs in J_p, the motion usually consists of translations and rotations, and the relative positions of J and J_p, (x_c, y_c, z_c), are related only to the rotation of J_p, which can be determined by the matrix of angular displacements M and the initial relative position (x_0, y_0, z_0) are calculated to obtain equation (2):

$$\begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} = M \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} \tag{2}$$

where the angular displacement matrix M consists of the three sub-rotation matrices of the node J_p around the ZXY axis connected sequentially as in equation (3):

$$M = RPY \tag{3}$$

The three subrotation matrices R, P, and Y are computed from the Roll angle r, the Pitch angle p, and the Yaw angle y in equations (4)-(6):

$$R = Rz(-r) = \begin{bmatrix} \cos(-r) & \sin(-r) & 0 \\ -\sin(-r) & \cos(-r) & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4}$$

$$P = Rx(-p) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(-p) & \sin(-p) \\ 0 & -\sin(-p) & \cos(-p) \end{bmatrix} \tag{5}$$

$$Y = Ry(-y) = \begin{bmatrix} \cos(-y) & 0 & -\sin(-y) \\ 0 & 1 & 0 \\ \sin(-y) & 0 & \cos(-y) \end{bmatrix} \tag{6}$$

Assuming that the 3D coordinates of J_p after that motion are (x_{pc}, y_{pc}, z_{pc}), then the 3D coordinates of J at the present moment (x_{Jc}, y_{Jc}, z_{Jc}) are equation (7):

$$(x_{J_c}, y_{J_c}, z_{J_c}) = (x_c, y_c, z_c) + (x_{pc}, y_{pc}, z_{pc}) \tag{7}$$

It can be seen that for the data recorded in the form of Euler angles, in order to calculate the 3D coordinates of a child node J, the 3D coordinates of its parent node J_p need to be calculated first.

In the human skeleton defined by the BVH motion capture data file, only the root node has its 3D coordinate information recorded in its motion data, while other non-root and non-leaf nodes only have their three rotation angles around the ZXZ axis recorded in their motion data. Therefore, determining the position of a particular non-root node requires sequentially determining the position of its predecessor nodes up to the root node. For a chain of nodes $J, J_1, J_2 \cdots J_r$ (where the parent node of leaf node J is J_1 and J_r is the root node of the chain of nodes), assuming that $M_1, M_2 \cdots M_r$ is the corresponding angular displacement matrix, then Eq. (2) can be be extended to Eq. (8):

$$\begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} = M_r \cdot \left(M_{r-1} \cdot \ldots \left(M_2 \cdot \left(M_1 \cdot \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} \right) \right) \right) \tag{8}$$

From the angular displacement matrix, the relative position, i.e., the offset, of a non-root node to its predecessor node on that node chain can be calculated, such that the offset of J with respect to J_1 is O_0, and such that the other offsets on the node chain are, in order, $O_1, O_2, \cdots O_{r-1}$, then Eq. (7) can be expanded to Eq. (9):

$$P_J = P_r + O_{r-1} + ... + O_2 + O_1 + O_0 \tag{9}$$

where P_J is the 3D coordinate of node J and P_r is the 3D coordinate of the root node J_r. As a result, the three-dimensional coordinates of any node in the node chain can be calculated, thus completing the data conversion from Euler angle form to three-dimensional coordinate form.

III. B. Construction of the model

III. B. 1) Network overview

Compared with other traditional deep learning methods, 3DCNNs are not limited to the input of 2D single-frame images, and can extract features from both temporal and spatial dimensions, and are able to extract motion information from multiple consecutive frames. Therefore, in this paper, 3DCNNs are used to recognize the skeleton information of typical movements of folk dances, which is single-channel information, with less computation and better model recognition performance. The 3D deep convolutional neural network used in this paper includes 4 convolutional layers, 2 downsampling layers, 2 fully connected layers and 1 Softmax classification layer. The downsampling layer uses Max-pooling with kernel size $3\times3\times3$ and step size 1.

III. B. 2) Detailed model structure

To capture motion information in multiple consecutive frames, features are computed in both spatial and temporal dimensions. The value of the cell with position coordinates (x, y, z) in the j th feature map of the i th layer is given by the formula shown in equation (10):

$$V_{ij}^{xyz} = f\left(b_{ij} + \sum_r \sum_{l=0}^{l_{i-1}} \sum_{m=0}^{m_{i-1}} \sum_{n=0}^{n_{i-1}} \omega_{ijr}^{lmn} v_{(i-1)r}^{(x+l)(y+m)(z+n)}\right) \tag{10}$$

where the 3D convolutional kernel has a time dimension of n_i and the position (l, m, n) is connected to the r th feature map the weight value of the convolutional kernel is ω_{ijr}^{lmn}.

The ReLU function is the most commonly used activation function in deep learning models. This function can make the parameters of the model sparse, thus reducing overfitting. In addition to this, it also reduces the amount of model computation.The ReLU activation function is defined as equation (11):

$$f(x) = \max(0, x) = \begin{cases} 0, & (x \le 0) \\ x, & (x > 0) \end{cases} \tag{11}$$

The calculation of maximum pooling in the model is shown in equation (12):

$$V_{x,y,z} = \max_{0\le i\le s_1, 0\le j\le s_2, 0\le k\le s_3} \left(\mu_{x\times s+i, y\times t+j, z\times r+k}\right) \tag{12}$$

where μ denotes the three-dimensional input vector, V denotes the output after the pooling operation, and s, t, r denotes the sampling step in the direction.The Softmax function is often used in the last layer of the classification task, which maps an n -dimensional vector x into a probability distribution such that the probability of the correct category tends to 1 and the others tend to 0, and such that the probability of all the categories sum to 1.

The phenomenon of overfitting is a phenomenon that performs poorly on test data when training a model, and the Dropout technique refers to temporarily discarding some neurons in the model according to a certain ratio when the deep learning model is iterated so that they do not participate in the model computation, thus significantly reducing the risk of overfitting. Dropout technique is also used in this paper to reduce the overfitting phenomenon.

The experiment uses a 5-fold cross-validation method, where the preprocessed dataset of typical movements of minority dances is randomly divided into 5 groups, of which 4 groups are used as the training set and 1 group is used as the test set, and then the results are averaged 5 times as the recognition rate.

IV. Testing of the model and evaluation of the application

IV. A. Segmentation of action sequences

In this section, the motion sequence segmentation experiment uses the Guanzhong gong dance recorded by the motion capture device as the data set, and selects the third group of dance movements with small misstep combinations as the sample data. The action video contains a total of 1125 frames, manual segmentation to remove the pre-preparatory action and the final ending action can be determined that there are four similar action combinations, at the beginning of the dance action there is a long period of time for the actor to prepare for the action, and there is the end of the action after the actor returned to the initial state of the process from the last position leading to the end position of the joints distance changes gently and then there is a drastic change.

From Table 1, it can be seen that the first segmentation position is at frame 110, between the preparatory action frame sequences, so there is no loss of action information. In the segmentation results of the other action sequences due to the existence of an obvious action convergence process in the middle of the small misstep and leg raising action, it is mistakenly recognized as a separate simple action from the 110th frame to the 280th frame, whereas under normal circumstances, the small misstep and the rear leg raising action belong to a dance action sequence. The results of comparing the manual segmentation position with the segmentation position determined by the modeling algorithm of this paper are shown in Table 1, and the first line in each segmentation position is the manual segmentation position.

Table 1: Manual segmentation and automatic segmentation of positions

Motion clip	Start frame	End frame	Textual	Motion description
1	0	130	111	Preparatory action
2	110	293	236	Lift your leg after taking a small wrong step
	280	353	366	Unrelated actions
3	340	423	503	Lift your leg after taking a small wrong step
	410	493		Unrelated actions
4	480	557	639	Lift your leg after taking a small wrong step
	544	611		Unrelated actions
5	598	680	712	Lift your leg after taking a small wrong step
	667	743	731	Unrelated actions

It can be found that using the double convention of distance change of the actor's joint coordinate position in neighboring frames and the number of frames per action sequence, the correct rate of manual segmentation can be increased a lot, the first action sequence in the sample is still segmented incorrectly because there is an obvious closing process between the small misstep and the rear leg lift action, and the action duration is longer. The other automatic segmentation positions are in irrelevant movements or have a small gap with manual segmentation, so they belong to the correct segmentation results. Since the initial segmentation position has more incorrect segmentation positions, the difference constraints between frame sequences and the difference constraints between neighboring minima and maxima are added to the model in this paper. The comparison results of segmentation positions after adding constraints are shown in Table 2.

Table 2: Comparison of segmentation positions with added constraint conditions

Motion clip	Start frame	End frame	Textual	Motion description
1	0	130	111	Preparatory action
2	110	293	236	Lift your leg after taking a small wrong step
	280	353	366	Unrelated actions
3	340	423	503	Lift your leg after taking a small wrong step
	410	493		Unrelated actions
4	480	557	639	Lift your leg after taking a small wrong step
	544	611		Unrelated actions
5	598	680	712	Lift your leg after taking a small wrong step
	667	743	731	Unrelated actions

After segmentation, the segmented action sequences are compared for similarity to determine the degree of similarity between the action sequences in the same video, thus determining the main action sequences that can represent the video. In this paper, we use the regularization distance for comparison, the smaller the cumulative regularization distance is the greater the degree of similarity between the action sequences. If six action sequences A0, A1, A2, A3, A4, A5 are segmented in a small staggered forward march, the distance between two and two action sequences is experimentally derived as shown in Table 3.

Table 3: Action sequence distance

Action sequence	A0	A1	A2	A3	A4	A5
A0	0	433.1067	2015.9431	1722.7724	2698.413	1309.4433
A1	674.6075	0	579.6003	475.9198	897.2526	346.2688
A2	3621.6946	647.2682	0	220.4263	55.8923	345.7317
A3	2885.9217	390.8994	80.1479	0	250.6787	345.6923
A4	4750.7619	1011.1404	85.6394	331.7224	0	536.2629
A5	1990.6734	426.1199	220.5923	278.2721	283.0742	0

Among the six action sequences, A0 is a preparatory action, which does not contain obvious action information, and it can be seen in the table that the regularization distances between A0 and the other action sequences are very large, except for the regularization distance with the A1 action which is 674.6075, and the regularization distances with the other actions are all in the range of 1800.0000 and above. The regularization distances between the remaining several action sequences are relatively small, and it can be determined that similar actions are generated in the remaining several action sequences. The first 10 minimum distances (5*2) can be selected to determine that the similar actions are generated in A2, A3, A4, A5, of which A2, A3, A4 are the most similar, and finally the longest action sequence A3 (340 frames~410 frames) is selected among several action sequences to represent the video information of the small staggered forward marching. The above 5*2 selection is based on the existence of all action sequences are similar (except the preparatory action), so it is necessary to ensure that the distance between all the action sequences are obtained.

IV. B. Action segmentation results

The human dance posture movement dataset was selected as the basis for this experimental study. The dataset consists of 50 samples of human dance movements with different postures and dance types, with a total number of 52,960 movement samples. The Guanzhong Drum and Gong Dance gesture action video of the dataset is selected, and the video is segmented into six different action segments, which are (FG-01) knee bending, (FG-02) big kick, (FG-03) turning circle, (FG-04) toe crunching, (FG-05) drum and gong dance step, and (FG-06) toe drawing circle, and the segmentation results are shown in Table 4.

Table 4: The segmentation result of dance movement videos

Fragment Number	Starting frame (pieces)	End frame (pieces)
FG-01	1	287
FG-02	288	346
FG-03	347	409
FG-04	410	468
FG-05	469	517
FG-06	518	646

IV. C. Human displacement localization results

Figure 2 shows the human displacement localization results of the model in this paper. As can be seen from Fig. 2, the human displacement localization data are basically stable under the range-of-view (RV) condition. The fluctuation ranges of X-axis and Z-axis under non-range-of-view (NRV) conditions are 127~138 and 127~143, respectively, and the overall fluctuation amplitude is small, which is basically the same as that under range-of-view conditions. This shows that the human displacement localization module modeled in this paper has good localization effect.

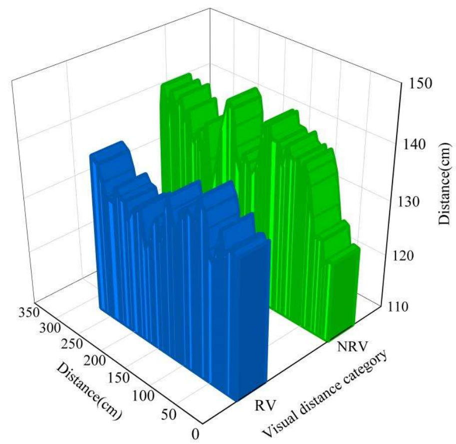

Figure 2: Human body displacement positioning results

IV. D. Comparison with similar algorithms

To further analyze the effectiveness and superiority of the dance movement recognition module of this paper's model, experiments were conducted to compare the recognition accuracy of this paper's model with that of the commonly used recognition algorithms KNN algorithm and the standard SVM algorithm, and the results are shown in Fig. 3.The selected recognition movements are (Y1) Pitching, (Y2) Tilting, (Y3) Punching, (Y4) Twisting, (Y5) Twisting, (Y6) Kicking, (Y7) Squatting, and (Y8) Jumping.

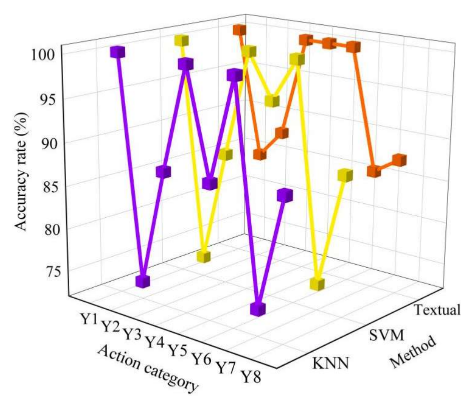

Figure 3: The results of dance movement recognition by different algorithms

As can be seen from the figure, compared with the KNN algorithm and the standard SVM algorithm, the model in this paper has the highest recognition accuracy for different dance movements, not only for the eight dance movements the recognition accuracy reaches 85% and above, and the recognition accuracy of multiple dance movements is as high as 100%. This shows that the model algorithm in this paper can recognize Guanzhong gong and drum dance movements more accurately and effectively, and has certain superiority.

V. Conclusion

In this paper, computer vision technology is integrated into the recognition of Guanzhong Drum and Gong Dance dance movements, and the action recognition model based on 3D CNNs algorithm is designed to accurately capture the action nodes, recognize the action categories, analyze the artistic features of the dance movements, and provide effective technical support for the construction of Guanzhong Drum and Gong Dance inheritance path. The designed action recognition model can effectively segment the action sequences and obtain the regular distance between the sequences. And it is more stable in human displacement localization, and the fluctuation range is controlled between 127~143. Compared with similar model algorithms, the action recognition model based on 3D CNNs algorithm shows excellent accuracy in recognizing Guanzhong gong and drum dance movements, with the recognition accuracy of 85.00% and above for all eight dance movements, and 100.00% for multiple dance movements.

Funding

This work was supported by Shaanxi Province Art Science Planning and Research Project, the project name is "Research on the Artistic Characteristics and Diversified Inheritance Path of "Gong Inspiration in Guanzhong" Based on Modern Popular Elements" (project number: SYE2024002).

References

[1] Xiao, L. (2022). Intangible cultural heritage reproduction and revitalization: value feedback, practice, and exploration based on the IPA model. Computational Intelligence and Neuroscience, 2022(1), 8411999.

[2] Tan, B. (2021). Thoughts on the Inheritance of Shaanxi Folk Dance in Universities. The Educational Review, USA, 5(7).

[3] Fan, T., & Xue, D. Q. (2018). Sustainable development of cultural industry in Shaanxi province of Northwest China: A SWOT and AHP analysis. Sustainability, 10(8), 2830.

[4] Wulf, C. (2020). Performativity and dynamics of intangible cultural heritage. In Ritual, Heritage and Identity (pp. 76-94). Routledge India.

[5] Zhang, Z. (2023). Research on the Protection and Inheritance Path of Intangible Cultural Heritage Dance. Highlights in Art and Design, 3(1), 50-53.

[6] Mitric, J., Radulovic, I., Popovic, T., Scekic, Z., & Tinaj, S. (2024, February). AI and Computer Vision in Cultural Heritage Preservation. In 2024 28th International Conference on Information Technology (IT) (pp. 1-4). IEEE.

[7] Deng, B., Sun, A., Shan, P., & Rojo Ortiz, C. A. (2024, December). Research on the Three-Dimensional Acquisition and Restoration Method of Cultural Relics Based on Computer Vision. In Proceedings of the 2024 2nd International Conference on Advances in Artificial Intelligence and Applications (pp. 180-184).

[8] Qi, Y., & Zhou, Q. (2024). Digital Protection and Inheritance of Intangible Cultural Heritage of Clothing Using Image Segmentation Algorithm. Computer-Aided Design and Applications, 21, 159-173.

[9] Ollaberganova, M. D., & Xudaybergenov, T. A. (2024). Exploring motion capture algorithms in computer vision using intel depth camera. Multidisciplinary Journal of Science and Technology, 4(10), 49-55.

[10] Park, J. W., & Park, S. M. (2024). Ownership Sharing System using Computer Vision-Based CNN Algorithm for Preserving National Heritage. The Journal of the Korea institute of electronic communication sciences, 19(6), 1347-1352.

[11] Chen, L., Wu, L., & Wan, J. (2025). Damage detection and digital reconstruction method for grotto murals based on YOLOv10. npj Heritage Science, 13(1), 91.

[12] Hu, Q., Yu, D., Wang, S., Fu, C., Ai, M., & Wang, W. (2017). Hybrid three-dimensional representation based on panoramic images and three-dimensional models for a virtual museum: Data collection, model, and visualization. Information Visualization, 16(2), 126-138.